Economics Education and Research Consortium

Working Paper Series

ISSN 1561-2422

No 04/10

Regional Reallocation of Russian Industry in Transition

Elizaveta Cheviakhova
Oleg Rytchkov

This project (02-086) was supported by the Economics Education and Research Consortium

Research area: **Enterprises and Product Markets**

JEL classification: R11, R12.

CHEVIAKHOVA E.Yu., RYTCHKOV O.A. Regional Reallocation of Russian Industry in Transition: A New Economic Geography Perspective.– Moscow: EERC, 2004.

In this paper we suggest to use a "new economic geography" paradigm for explanation of regional reallocation of industrial employment in Russia in 1985-1999. We construct a "new economic geography" type model adjusted to specific features of Russian economy. This model gives a counterfactual distribution of industry across regions and allows us to construct a theoretical factor $NEGF$ which is supposed to predict real changes in allocation of industrial employment. Our analysis of empirical data shows that $NEGF$ indeed has a predictive power and this result is valid for a sufficiently wide range of model specifications.

Acknowledgements. We are grateful to Michael Alexeev, Richard Ericson, Gary Krueger and Mark Schaffer for numerous valuable comments and suggestions. We also appreciate helpful discussions with Ruben Enikolopov, Sergei Guriev and Ksenia Yudaeva.

Keywords: Russia, new economic geography, industry allocation

Elizaveta Cheviakhova
Boston College, Department of Economics,
140 Commonwealth Avenue, Chestnut Hill, MA 02467-3806, USA
Tel. +7(101)617 552 36 70
Fax +7(101)617 552 23 08
E-mail: cheviakh@bc.edu

Oleg Rytchkov
MIT, Sloan School of Management,
50 Memorial Drive, Cambridge, MA 02142, USA
Tel. +7(101)617 253 36 37
E-mail: rytchkov@mit.edu

Contents

Non-technical summary

It is a widely recognized fact that economic activity in general and industrial production in particular are unevenly distributed in space. One of theoretical explanations of non-uniform distribution of industrial production is given by a strand of economic literature tabbed as "new economic geography". Basically, "new economic geography" considers the existing geographic structure of industry as a result of individual choices made by firms which select their locations facing a tradeoff between increasing returns to scale and transportation costs.

In general, testing predictions of "new economic geography" is not an easy task because of a variety of other factors which affect location choices of firms. The basic idea of our research is to test implications of "new economic geography" by analyzing regional reallocation of Russian industry in the transition period. Indeed, transition economies have experienced a significant industry restructuring which could not help altering the spatial allocation of industries. This restructuring was the result of invoking market forces (captured by "new economic geography" models) which did not play any significant role in the Soviet period. Thus, if "new economic geography" paradigm is valid it should be capable to explain the observed changes in spatial pattern of industrial production in Russia.

To implement this idea we construct a model of a three-sector economy which has basic features of "new economic geography" models and accounts for specific Russian conditions (by latter we understand the observed low labor force migration in Russia and rapid reallocation of resources between sectors). This model gives a counterfactual equilibrium allocation of industry understood as a distribution of industrial employment across regions. The difference between this equilibrium allocation and the allocation observed at the beginning of transition should be used as a predictor of observed changes in allocation of industry.

Practically, it appears that the system on non-linear equations determining the equilibrium allocation is extremely complicated and can have multiple solutions. Instead of deriving equilibrium allocation explicitly we follow a round-about way. Under assumption that initial allocation is not far from the equilibrium one, we can find the difference between the initial allocation and the equilibrium in linear approximation. We call this vector of differences New Economic Geography Factor ($NEGF$) and by construction it represents the "new economic geography" predictions of the industry employment reallocation.

Main results of our research come from the empirical part. Regression analysis shows that the constructed factor indeed has a substantial predictive power and explains about 15% of total variation. Moreover, this result is robust to a number of changes of model specifications such as the form of functional relation between transportation costs and distances between regions as well as values of directly unobserved parameters. Also the result is not sensitive to the choice of the final year. To separate the effect of the new economic geography factor from other potential explanations we use a number of control variables in our regressions. Notably they do not spoil the predictive power of $NEGF$ and most of them appear to be insignificant. The only significant one is the export dummy which separates export-oriented and import-oriented regions. Overall, the main conclusion of the paper is that "new economic geography" indeed works and gives an explanation of regional reallocation of Russian industry in 1985 - 1995.

1 Introduction

It is a widely recognized fact that economic activity in general and industrial production in particular are unevenly distributed in space. One of theoretical explanations of non-uniform industrial production distribution is given by a strand of economic literature tabbed as "new economic geography". Basically, this approach hinges on various linkages emerging in economy with positive transportation cost and increasing returns to scale production technology. In particular, due to increasing returns to scale firms tend to locate all their production in one place. However, since transportation costs are non-zero a firm which places all activity in one location has to pay large transportation costs for delivering its output to geographically dispersed customers. According to the viewpoint of "new economic geography" this tradeoff is a major determinant of spatial distribution of industrial production.

Technically, most of "new economic geography" models are based on the Dixit-Stiglitz model of monopolistic competition (Dixit and Stiglitz, 1977). Krugman (1991) suggested to incorporate the ideas of Dixit and Stiglitz into spatial setting and this allowed him to construct the first general equilibrium model of spatial distribution of industrial production. In his paper Krugman shows how an economy may become differentiated into "manufacturing" and "agricultural" geographical clusters. Namely, in his model the centripetal effect acts as follows: the larger number of firms choose the same location, the more labor they need as input and, hence, the larger local market for their products becomes (it is assumed that workers can move between regions). Under non-zero transportation cost, this effect draws all manufacturing firms to one location. But, as agricultural workers are assumed to be immobile in that model, the other location never becomes completely depopulated. Depending on the size of transportation cost, centripetal or centrifugal force prevails. Thus, under certain (intermediate) values of transportation cost a core-periphery economy with predominantly manufacturing and agricultural regions emerges.

While in Krugman (1991) a self-reinforcing agglomeration process relies on the fact that the larger labor market means the larger market for final product, in Venables (1996) agglomeration originates from internal (demand and cost) linkages within the industry. Interaction between final product and intermediate industries (both monopolistically competitive, producing differentiated product) leads to agglomeration of industry if transportation cost is sufficiently low. These two papers represent two major approaches to explanation of agglomeration within a new economic geography strand of literature. Also these two approaches were used to address questions related to economic growth (e.g., Baldwin, 1999; Baldwin and Forslid, 2000; Martin and Ottaviano, 2001; Puga and Venables, 1999), international trade and development (Krugman and Venables, 1995).

Most of the new economic geography literature exploit such modelling elements as Dixit-Stiglitz monopolistic competition, "iceberg type" transportation costs, and Cobb-Douglas specification of preferences. In contrast, Ottaviano et al. (2002) use an alternative specification: quadratic utility function and transportation cost in terms of a numeraire good. Still, they also obtain the result similar to Krugman (1991) that degree of agglomeration depends on transportation cost and this relationship has a U-shape form.

Empirical research in "new economic geography" is much more scarce than the theoretical one. Indeed, citing Fujita et al. (1999) "economic models with increasing returns and imperfect competition have proved difficult subjects for empirical work". Despite this there are few papers that attempt to test the validity of the main theoretical results. Davis and Weinstein (1999) analyze allocation of industries in Japan and show the existence of the economic geography effects in eight of nineteen manufacturing sectors. Making use of US county data Hanson (1998) tests the significance of the parameters in the Krugman (1991) model and basically confirms the relevance of the model and its main predictions. The similar analysis for German city-districts is performed in the paper by Brakman et al. (2001) which also confirms the predictions of the model. However, these two papers concentrate mostly on the new economic geography implications for the distribution of wages, but not for the allocation of industries. Combes and Lafourcade (2001) estimate a structural model of the NEG type for France and use it to investigate the decline of transportation cost as a cause of regional inequality. They find evidence that intermediate inputs and geographical features are important determinants of the concentration pattern of French economic activities.

Also, new economic geography models were extensively used for research in international trade. One of the main questions addressed by this literature is how the fall of trade barriers influences location of production and welfare of different nations. For example, Midelfart-Knarvik et al. (2000) estimate the location of production in the EU and show that endowments of skilled labor as well as forward and backward linkages within industry are important determinants of industrial structure.

The basic idea of our research is to apply the “new economic geography” framework for explanation of industrial reallocation in Russia in the transition period. The benefits of this approach are twofold. On one hand, “new economic geography” can yield a solid theoretical explanation of this extremely complicated process and, thus, a deeper understanding of the underlying economic forces which drive industry reallocation. On the other hand, the examination of rapidly evolving transition economy can provide a powerful test for the theory itself. This focus on dynamical aspects is the basic distinction of our approach from the most of the empirical “new economic geography” literature which mainly analyze the existing geographical structure of industry and explains the agglomeration phenomenon whereas in our research we concentrate on the determinants of industry reallocation dynamics.

There are several arguments why the “new economic geography” framework could be relevant for studying transition economies and why it is reasonable to suggest that “new economic geography” factors should have played an important (and probably predominant) role in the geographical reallocation of industry.

First, it is reasonable to expect that in transition economies the speed of industry reallocation is much greater than in most developed economies. Indeed, transition economies have experienced a significant industry restructuring which could not help altering the spatial allocation of industries. For example, in the 1990s the Russian economy has been undergoing a process of fundamental institutional and structural transformation. While in 1991 the share of service sector in GNP was only about one quarter, by 1995 it increased more than twofold. This rise was achieved at the expense of industrial production and in view of abrupt decline of GNP. And this is only one piece of evidence of massive reallocation of resources in the economy and this reallocation potentially had its geographical dimension as well.

Next, “new economic geography” factors are essentially based on market forces which did not play any role in the Soviet period but started to have an effect in the period of transition. Indeed, in the “new economic geography” paradigm the geographical structure of industry is a result of independent decisions taken by the economic agents (workers and firms). In Soviet period the allocation of plants was governed by the central planning agency and might be substantially different from the allocation that firms would have taken if they had decided themselves. After elimination of the planning economy firms got this right and they might have started the reallocation of resources as a result of their profit maximization. From a theoretical point of view this process can be considered as a transition from a non-equilibrium state (which is characterized in particular by the industry allocation pattern) to the equilibrium one. Since the market forces, which are responsible for the geographical structure of industry in market economies, switched on almost instantaneously they became more clearly revealed than in developed countries where reallocation processes are much slower. This means that transition economies can provide a unique natural laboratory for investigation of dynamics of industry reallocation and its determinants. In some sense, these factors work actively before our eyes and this substantially simplifies the task of their identification.

In our study we mostly consider the industry as a whole without concentration on specific branches. Indeed, when we look at the aggregated data there is a hope that some factors which are specific for particular industries will play no role on average and it will be possible to identify factors, which are common to at least most industries. It is reasonable to believe that in this way "new economic geography" factors will be elicited since they are supposed to act uniformly and to be responsible for industry reallocation.

From the beginning we need a well-defined characteristic of industry allocation. In general, one can base this characteristic on different variables. First of all, the industry allocation can be described by shares of industrial output produced in the given region. However, this approach has several drawbacks. Aggregated output can be calculated only in terms of prices, but not in terms of real units. This introduces significant problems since output prices vary among industries but industries are not allocated uniformly across the regions, each region has its own profile of industrial output. Consequently, the characteristic of industry allocation constructed on the basis of aggregate industrial output will be essentially a measure of not only the industrial production in the given region but also the industrial profile. Moreover, prices of different goods do not change strictly proportionally. As a result, one should use regional price deflators to compare shares of industrial output in different years. All this makes the measure of industry allocation based on industrial output very noisy and poorly defined.

To circumvent the discussed problems we construct a characteristic of industry allocation making use of industrial employment, namely, we measure the presence of industry in the region by the regional share of industrial employment. On one hand, this measure is well defined and less noisy than that based on output. On the other hand, this is in accordance with the existing practice in the economic geography literature.

Methodologically the strategy of testing the "new economic geography" explanation of industry reallocation in Russia is quite simple: take a "new economic geography" type model, find a counterfactual equilibrium allocation of industry understood as a distribution of industrial employment across regions, compare it with the initial allocation and using empirical data check that this difference is a predictor of observed changes. Unfortunately, this way is not as straightforward as it could be. Below we discuss several complications that should be overcome.

First of all, we need a model which on one hand inherits basic ideas of "new economic geography" and on the other hand is adjusted to specific Russian conditions. By latter we understand the observed low labor force migration in Russia and rapid reallocation of resources between sectors. Indeed, if there are high barriers for labor migration then the migration-induced linkages which are the key factor determining the geographic structure of industry in Krugman (1991) do not work. Also it is crucial for our purposes that this model predicts smooth distribution of industrial production among the regions (note, that this is not the case in the standard Krugman model, in which only two extremes can realize: production is uniformly allocated across regions or is concentrated only in one region).

Next complication follows from the fact that the equations describing the model equilibrium are extremely complicated and in general we are unable to solve them even numerically. To circumvent this problem we need additional assumption that the initial allocation is not far from the equilibrium one. Under this condition we can find a difference between the initial allocation and the equilibrium in the linear approximation. We call this vector of differences New Economic Geography Factor ($NEGF$) and by construction it represents the "new economic geography" predictions of industrial employment reallocation.

Main results of our research come from the empirical part. First of all, the regression analysis shows that $NEGF$ indeed has a substantial predictive power and explains about 15% of total variation. Moreover, this result is not sensitive to a number of changes of model specifications such as the form of the functional relation between transportation costs and the distances between regions as well as values of directly unobserved parameters. Also the result is robust to the choice of the final year. To separate the effect of the new economic geography factor from other potential explanations we use a number of control variables in our regressions. Notably they do not spoil the predictive power of $NEGF$ and most of them appear to be insignificant. The only significant one is the export dummy which separates export-oriented and import-oriented regions.

To be sure that we try to explain a trend in the industry allocation but not random fluctuations we supplement our regression analysis with the test of a null hypothesis that there was no trend in industry reallocation across regions. To this end we elaborate a special procedure to overcome the problem arising from a small number of time observations. Ultimately we find that data are not compatible with the hypothesis that there was no trend in the reallocation process.

2 Model

In this Section we construct a model of a "new economic geography" type, which is purposely designed to incorporate important stylized characteristics of Russian economy. First, this is a three sector model (there are industry sector, agricultural sector and service sector). The purpose of this structure is to capture the effects of rapid reallocation of resources between sectors and, in particular, rapid growth of the service sector. Second, we assume that workers are geographically immobile, but can move between sectors. This assumption is in accordance with the observed low labor force migration in Russia (Andrienko and Guriev, 2004), but significant shifts between sectors. Combination of immobile labor force and non-tradable services allows us to make not only qualitative predictions, but also quantitative ones. Indeed, in this setup we obtain equilibria with non-trivial distribution of industry employment across regions. Also this distribution smoothly depends on the parameters of the model and it may allow us to perform comparative statics. To preserve the generality of consideration, we construct a model with arbitrary number of regions R. In our empirical application we will put $R = 78$. The basic notations of the model are collected in Table 1.

2.1 Production side

Consider a world consisting of R regions. We assume that economy contains three sectors in each region: industry (M), agriculture (A) and services (S). These sectors differ with respect to the type of output (homogeneous or differentiated), production function (CRS or increasing returns to scale), market structure and transportation costs. Output of these sectors will be denoted as q^m, q^a and q^s respectively. Also we assume that there is only one production factor which is labor. This makes the model tractable and consistent with our description of industry allocation by means of shares of industrial employment.

Characteristic properties of the sectors are summarized in Table 2. In particular, the industrial sector produces a variety of differentiated goods, the number of varieties in each region is $n_r,\ r = 1..R$. Technology is the same for each variety and involves a fixed input F and a marginal input requirement c. Namely, it takes $l^m = F + cq^m$ labor units to produce the quantity q^m. Thus, production the in industrial sector is subject to increasing returns to scale. This property is common to most "new economic geography" models, which key feature is a tradeoff between increasing returns to scale and transportation costs. The industrial sector is monopolistically competitive and firms enter the market until they earn negative profit. The condition that each firm earns zero profit determines an equilibrium number of firms in each region.

As agricultural we define a sector where production needs no start-up investments in comparison with the industrial one, but where the returns to labor are also smaller. We also assume that the agricultural sector produces differentiated good, each region has its own type of agricultural output (in most "new economic geography" models the agricultural sector is supposed to be homogeneous). On one hand this assumption makes the model a bit more realistic and on the other hand allows to avoid several technical problems[1]. In each region the agricultural sector is competitive and has an access to a simple CRS technology $q^a = l^a$.

The service sector has the simplest structure in our model. We assume that it produces a homogeneous good and has an access to a linear technology $q^s = l^s$.

The regions under consideration are assumed to be geographically separated, i.e. there are non-zero costs of product transportation between them. The transportation costs for agricultural and manufactured goods are assumed to be finite and equal to $T_{ij} > 1$. It means that if a unit of a good is shipped from region i to region j, then only a fraction $1/T_{ij}$ of the original unit eventually arrives (this is a so called "iceberg form" of transportation costs, see Samuelson, 1954). It is natural to consider services as non-tradable goods and to put their transportation costs equal to infinity.

2.2 Consumption side

Now consider the consumption side of our model. We assume that there is fixed population (number of identical workers) L_r in each region r, i.e. workers can move only between sectors but not between regions. A consumer in each region has a Cobb-Douglas utility function

$$U = M^{\mu_m} A^{\mu_a} S^{\mu_s}, \qquad \mu_m + \mu_a + \mu_s = 1,$$

where S stays for services, A and M are composite agricultural and manufactured goods respectively:

$$M^{\rho_m} = \int_0^n m_i^{\rho_m} di, \qquad A^{\rho_a} = \sum_{i=1}^{R} a_i^{\rho_a}, \qquad 0 < \rho_m < 1, \qquad 0 < \rho_a < 1.$$

m_i and a_i denote consumption of each available variety and n is a number of manufactured varieties. The consumer in each region solves the following optimization problem

$$\max_{m_i, a_i, S} \left[\int_0^n m_i^{\rho_m} di \right]^{\frac{\mu_m}{\rho_m}} \left[\sum_{i=1}^{R} a_i^{\rho_a} \right]^{\frac{\mu_a}{\rho_a}} S^{\mu_s} \qquad \text{s.t.} \qquad \int_0^n p_i^m m_i di + \sum_{i=1}^{R} p_i^a a_i + p^s S = Y,$$

[1] In particular, if the agricultural sector produces a homogeneous good and transportation of this good is costly consumers in the given region in general will consume the good produced in a small number of regions and we will immediately have a "boundary type" solution.

where p_i^m, p_i^a and p^s are prices for manufacturing, agricultural and services respectively. Y is a regional income. This problem can be easily solved and the solution has the form

$$m_i = \frac{\mu_m Y}{(P^m)^{\rho_m/(\rho_m-1)}}(p_i^m)^{1/(\rho_m-1)} = \frac{\mu_m Y}{(P^m)^{1-\sigma_m}}(p_i^m)^{-\sigma_m}, \tag{1}$$

$$a_i = \frac{\mu_a Y}{(P^a)^{\rho_a/(\rho_a-1)}}(p_i^a)^{1/(\rho_a-1)} = \frac{\mu_a Y}{(P^a)^{1-\sigma_a}}(p_i^a)^{-\sigma_a}, \tag{2}$$

$$S = \frac{\mu_s Y}{p^s},$$

where P^a and P^m are composite price indices, which are equal to

$$P^m = \left[\int_0^n (p_i^m)^{\frac{\rho_m}{\rho_m-1}} di\right]^{\frac{\rho_m-1}{\rho_m}} = \left[\int_0^n (p_i^m)^{1-\sigma_m} di\right]^{\frac{1}{1-\sigma_m}}, \tag{3}$$

$$P^a = \left[\sum_{i=1}^R (p_i^a)^{\frac{\rho_a}{\rho_a-1}}\right]^{\frac{\rho_a-1}{\rho_a}} = \left[\sum_{i=1}^R (p_i^a)^{1-\sigma_a}\right]^{\frac{1}{1-\sigma_a}}. \tag{4}$$

Here we have introduced the following notation: $\sigma_m = 1/(1-\rho_m),\ \sigma_a = 1/(1-\rho_a)$.

The above formulas are valid for all regions. If we consider a particular region s we get that most goods consumed there are brought from other locations and, consequently, prices of such goods include transportation costs. Namely, the price of the good brought from the region r to the region s is

$$p_{rs}^m = p_r^m T_{rs}, \qquad p_{rs}^a = p_r^a T_{rs},$$

where p_r^m and p_r^a are prices of particular manufactured and agricultural good in the region r. To simplify notation we suppress the index pertaining to a particular good. Without losing much generality we assume that each region has both manufactural and agricultural sectors. Taking this into account we obtain aggregate demand from (1) and (2). Namely, demand for one of the manufactured products from the region r has the form

$$q_r^m = \sum_{s=1}^R m_{rs} T_{rs} = \sum_{s=1}^R \frac{\mu_m Y_s}{(P_s^m)^{1-\sigma_m}}(p_r^m T_{rs})^{-\sigma_m} T_{rs}. \tag{5}$$

Similarly, demand for the agricultural product from the region r is

$$q_r^a = \sum_{s=1}^R a_{rs} T_{rs} = \sum_{s=1}^R \frac{\mu_a Y_s}{(P_s^a)^{1-\sigma_a}}(p_r^a T_{rs})^{-\sigma_a} T_{rs}. \tag{6}$$

2.3 Equilibrium

Now we combine the consumption and production sides of our economy. The assumption that workers can freely move between sectors implies that in equilibrium wages are the same in all sectors. However, immobility of workers across regions prevents equalizing of wages and we denote the wage in region r as w_r. Consequently, aggregate income in the region r is $Y_r = w_r L_r$, where L_r is total population in the region r. This formula implicitly captures that in equilibrium profits of manufacturing firms are zero. Since the agricultural and service sectors are competitive and use simple linear technology we have that in equilibrium $p_r^a = p_r^s = w_r$.

Industrial sector is monopolistically competitive. It means that each incumbent manufacturing firm is a monopolist on the market of its own good and solves the following profit maximization problem:

$$\pi_r = p_r^m q_r^m - w_r(F + cq_r^m) \to \max_{p_r^m} \tag{7}$$

where q_r^m is demand given by (5) for a particular good produced in the region r. Substituting (5) into (7) and performing maximization with respect to p_r^m we get the following solution:

$$p_r^m = \frac{cw_r}{\rho_m}.$$

Note that here we implicitly assumed that the contribution of the price of particular good into the price index is negligible. If the incumbent firm sets this price, its profit will be equal to

$$\pi_r = \frac{cw_r q_r^m}{\sigma_m - 1} - w_r F. \tag{8}$$

In equilibrium firms do not have incentives to enter or exit the market and this can be valid only if their profits are equal to zero: $\pi_r = 0$. This gives an equilibrium output of each firm:

$$q_r^m = \frac{F(\sigma_m - 1)}{c}.$$

Consequently, employment of each industrial firm is $F + F(\sigma_m - 1) = F\sigma_m$. Denote the number of industrial firms (the number of varieties) in the region r as n_r. Then the total industrial employment in the region r is $l_r^m = n_r F \sigma_m$.

To determine the equilibrium values of w_r, n_r, p_r^m we consider a set of balance conditions. Equating supply of and demand for a particular manufacturing good we get

$$\frac{F(\sigma_m - 1)}{c} = \sum_{s=1}^{R} \frac{\mu_m w_s L_s}{(P_s^m)^{1-\sigma_m}} \left(\frac{c w_r}{\rho_m} T_{rs} \right)^{-\sigma_m} T_{rs}. \tag{9}$$

This is one of our basic equations. Further, the balance of agricultural good gives employment in the agricultural sector:

$$l_r^a = q_r^a = \sum_{s=1}^{R} \frac{\mu_a Y_s}{(P_s^a)^{1-\sigma_a}} (p_r^a T_{rs})^{-\sigma_a} T_{rs}.$$

Similarly, we get employment in the service sector:

$$l_r^s = q_r^s = \frac{\mu_s Y_r}{w_r} = \mu_s L_r.$$

The balance of labor $l_r^m + l_r^a + l_r^s = L_r$ gives our second basic equation

$$n_r F \sigma_m + \sum_{s=1}^{R} \frac{\mu_a w_s L_s}{(P_s^a)^{1-\sigma_a}} (w_r T_{rs})^{-\sigma_a} T_{rs} = L_r(1 - \mu_s). \tag{10}$$

As a last step, consider the price indexes P_s^a and P_s^m. From (3) and (4) we get

$$P_s^m = \left[\sum_{r=1}^{R} n_r (p_r^m T_{rs})^{1-\sigma_m} \right]^{\frac{1}{1-\sigma_m}} = \left[\sum_{r=1}^{R} n_r (\frac{c w_r}{\rho_m} T_{rs})^{1-\sigma_m} \right]^{\frac{1}{1-\sigma_m}},$$

$$P_s^a = \left[\sum_{r=1}^{R} (p_r^a T_{rs})^{1-\sigma_a} \right]^{\frac{1}{1-\sigma_a}} = \left[\sum_{r=1}^{R} (w_r T_{rs})^{1-\sigma_a} \right]^{\frac{1}{1-\sigma_a}}. \tag{11}$$

Finally, the substitution of (11) into (9) and (10) and rescaling of the constant F yield the following set of equations which determine the equilibrium values of n_r and w_r:

$$F n_r - L_r + \frac{\mu_a}{\mu_m} \left[\sum_{s=1}^{R} \frac{w_s}{w_r} \frac{(w_r T_{rs})^{1-\sigma_a}}{\sum_{q=1}^{R} (w_q T_{qs})^{1-\sigma_a}} L_s - L_r \right] = 0, \tag{12}$$

$$\sum_{s=1}^{R} \frac{w_s}{w_r} \frac{(w_r T_{rs})^{1-\sigma_m}}{\sum_{q=1}^{R} F n_q (w_q T_{qs})^{1-\sigma_m}} L_s - 1 = 0. \tag{13}$$

We have formulated the equilibrium conditions in terms of the variables n_r and w_r. However, since in the equilibrium the values of n_r are proportional to l_r^m and we work with the shares of industrial employment but not with its absolute values, two sets of variables (n_r, w_r) and (l_r^m, w_r) are absolutely equivalent.[2]

Scrutinizing equations (12) and (13) we can make several observations. First of all, note that the constants F and c dropped out, i. e. in our model the technological parameters are irrelevant to the regional industry shares and the distribution of wages. Indeed, the constant c even does not appear in the equations and the constant F enters only in combination Fn_i. This means that if the constant F is not specified the equations (12) and (13) allow to determine n_i only up to a scalar factor. However, this factor is actually irrelevant, since we are interested in shares but not in absolute values of n_i and in calculation of shares this factor will also drop out. Next, the preference parameters μ_a and μ_m enter in the equation only as a ratio. This means that only this ratio determines industry allocation. In particular, if μ_s goes up keeping the ratio μ_a/μ_m constant, this will not affect the equilibrium allocation of production. Thus it follows from our model that growth of the service sector (which corresponds to exogenous change in the preference parameter μ_s in the model) will not lead to reallocation of production despite the outflow of workers from the manufacturing sector if it does not change mutual preferences between industrial and agricultural goods.

2.4 New Economic Geography Factor (*NEGF*)

In Introduction we argue that regional reallocation of industry can be determined by market forces which make enterprizes adjust their activity to demand for their output and this may change the pattern of industrial allocation inherited from Soviet period. To test the role of these forces empirically we use our model and the idea of the test is quite simple. Equations (12) and (13) determine a counterfactual allocation of industry that would emerge if only "new economic geography" forces acted. First we compare this counterfactual allocation and the real allocation of industries at the beginning of transition period. Definitely, they will not coincide and we can deduce the difference between these

[2] Formally, we can work with the absolute values of industrial employment instead of its regional shares. However, the former measure is noisier and is influenced by a variety of other factors through the total industrial employment which are not relevant for answering questions about the regional distribution of industry.

allocations. This difference will be further referred to as theoretical. Unfortunately, we cannot directly calculate the counterfactual allocation, so to find the theoretical difference we use Taylor expansion. Afterwards we consider actual dynamics of industry allocation and construct the difference between real allocations in two different years (they should be separated by appropriate period during which the change of allocations becomes observable). Comparing theoretical differences with the real ones we can conclude whether the real reallocation of industries was at least partially driven by new economic geography factors or not.

There are two substantial problems pertaining to direct handling of counterfactual allocation given by equations (12) and (13). The first one is that these equations are extremely complicated (indeed, given 78 Russian regions we have 156 nonlinear equations with 156 unknowns) and it is a very hard computational problem to solve them even numerically. The second problem is that due to nonlinearity these equations can have multiple solutions and it could be very difficult to choose that specific equilibrium to which this system will converge starting from the given configuration. To overcome these problems we use a trick. First of all, we assume that the difference between the initial point and the target equilibrium is small. As a result, it is reasonable to find the difference between these allocations only in the linear approximation. To illustrate this approach denote the set of unknowns as a vector x, $x = \{n, w\}$ and consider the left hand side of the system as a vector-valued function $H(x)$. Thus our system of equations is $H(x) = 0$. Let x_0 be the initial allocation which in general does not solve the system and the substitution of it into equations gives a vector H_0 instead of 0: $H(x_0) = H_0$. Let $\bar{x}$ be an equilibrium: $H(\bar{x}) = 0$. We are interested in the difference $\bar{x} - x_0$ as this difference is supposed to explain the real dynamics. $H(x)$ is differentiable at x_0 so making use of the Taylor expansion of $H(x)$ at the point x_0 and taking $H(x)$ at the point $\bar{x}$ we get that in linear approximation

$$H(x_0) + H'(x_0)(\bar{x} - x_0) = 0,$$

where $H'(x_0)$ is a shortcut for the Jacobian of the system at the point x_0. Thus the vector $\bar{x} - x_0$ is

$$\bar{x} - x_0 = -H'(x_0)^{-1} H_0. \tag{14}$$

This approach solves simultaneously both problems discussed above. Indeed, matrix inversion is now the most computationally intensive operation, which is incomparably simpler than the procedure of solving the whole system. Moreover, this linearization procedure automatically chooses the equilibrium which is likely to be the closest one to the starting allocation. As a result, there is no problem of selecting appropriate equilibrium. Thus, we employ the linearization for constructing the vector of differences $\bar{x} - x_0$, which we will use as a predictor in the regressions. To be more concise, we will denote it in regressions as $NEGF$ (New Economic Geography Factor).

Theoretically the described linearization procedure is pretty straightforward. However, there are a number of subtleties in its practical realization. These subtleties will be discussed and treated one by one.

First of all, the system of equations (12) and (13) is invariant under rescaling of wages w, i. e if (n, w) is a solution than $(n, \lambda w)$, where $\lambda > 0$ is also a solution. In other words, wages are determined only up to a scalar factor and one of equations is redundant. It means, that for making the linearization applicable we have to fix the wage in one of the regions (otherwise the Jacobian matrix of the system is degenerate and cannot be inverted). To do this we put the wage in one of the regions equal to 1 (in our empirical work we take the Belgorod region as a benchmark). Note that after fixing the wage we get $2R - 1$ unknowns instead of $2R$ and for solving for them we do not need one of the system equations. Without loosing generality we discard the last one from (13).

Next complication related to the suggested linearization originates from the fact that we do not know actual values of n_i, but can operate only with their shares $s_i = n_i / \sum_{r=1}^{R} n_r$. It is worth to remind that when we considered the theory we attracted attention to the fact that unknowns n_i go together with undetermined constant F and actually we have equations not for n_i themselves but for the combinations Fn_i. However, it does not influence the final result which is stated in terms of shares, but not in terms of the real values n_i. To calibrate the system (to be able to substitute shares s_i instead of n_i) we should choose the constant F appropriately. Since performing the linearization procedure we should be able to substitute shares of employment from x_0 it is convenient to normalize F so that by definition if the point x_0 were a true solution then n_i would be equal to shares, i.e. $\sum_{r=1}^{R} n_r = 1$. This leads to the following definition of F:

$$F = \sum_{r=1}^{R} \left(L_r - \frac{\mu_a}{\mu_m} \left[\sum_{s=1}^{R} \frac{w_s}{w_r} \frac{(w_r T_{rs})^{1-\sigma_a}}{\sum_{q=1}^{R} (w_q T_{qs})^{1-\sigma_a}} L_s - L_r \right] \right),$$

where the wages w are taken at the point x_0. Obviously, this definition implies that in the true equilibrium $\bar{x}$ the sum of n_i is not equal to 1. Consequently, we should translate the differences $\bar{n}_i - n_{i0}$ obtained from $\bar{x} - x_0$ into differences of shares of employment. In the linear approximation this can be done by the following transformation:

$$\Delta s_i = \Delta \frac{n_i}{\sum_{r=1}^{R} n_r} = \Delta n_i - n_i \Delta n. \tag{15}$$

Now let us present a detailed calculation of the Jacobian matrix $H'(n, w)$. Actually we have

$$H(n, w) = \begin{pmatrix} H_1(n, w) \\ H_2(n, w) \end{pmatrix},$$

where $H_1(n, w)$ and $H_2(n, w)$ are left hand sides of equations (12) and (13) respectively. Thus the Jacobian matrix $H'(n, w)$ can be put as

$$H'(n, w) = \begin{pmatrix} \frac{\partial H_1}{\partial n}(n, w) & \frac{\partial H_1}{\partial w}(n, w) \\ \frac{\partial H_2}{\partial n}(n, w) & \frac{\partial H_2}{\partial w}(n, w) \end{pmatrix}, \tag{16}$$

where the constituent blocks are given by

$$\frac{\partial H_{1r}}{\partial n_u}(n, w) = F \delta_{ru}, \quad r = 1 \ldots R,\ u = 1 \ldots R,$$

$$\frac{\partial H_{1r}}{\partial w_u}(n, w) = \frac{\mu_a}{\mu_m} \sum_{s=1}^{R} \frac{L_s T_{rs}^{1-\sigma_a} w_r^{-\sigma_a}}{\sum_{q=1}^{R} (w_q T_{qs})^{1-\sigma_a}} \left[\delta_{su} - \sigma_a \delta_{ru} \frac{w_s}{w_r} - \frac{(1-\sigma_a) w_s w_u^{-\sigma_a} T_{us}^{1-\sigma_a}}{\sum_{q=1}^{R} (w_q T_{qs})^{1-\sigma_a}} \right], \tag{17}$$
$$r = 1 \ldots R,\ u = 2 \ldots R,$$

$$\frac{\partial H_{2r}}{\partial n_u}(n, w) = -\sum_{s=1}^{R} \frac{w_s}{w_r} \frac{L_s (w_r T_{rs})^{1-\sigma_m} (w_u T_{us})^{1-\sigma_m}}{F \left(\sum_{q=1}^{R} n_q (w_q T_{qs})^{1-\sigma_m} \right)^2}, \quad r = 1 \ldots R-1, u = 1 \ldots R,$$

$$\frac{\partial H_{2r}}{\partial w_u}(n, w) = \sum_{s=1}^{R} \frac{L_s T_{rs}^{1-\sigma_m} w_r^{-\sigma_m}}{\sum_{q=1}^{R} F n_q (w_q T_{qs})^{1-\sigma_m}} \left[\delta_{su} - \sigma_m \delta_{ru} \frac{w_s}{w_r} - \frac{(1-\sigma_m) w_s w_u^{-\sigma_m} T_{us}^{1-\sigma_m} n_u}{\sum_{q=1}^{R} n_q (w_q T_{qs})^{1-\sigma_m}} \right],$$
$$r = 1 \ldots R-1,\ u = 2 \ldots R.$$

Here δ_{ij} is Kronecker symbol: $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. For the normalization discussed above we have fixed $w_1 = 1$ and dropped the last equation from $H_2(n, w)$. The

Jacobian (16) should be calculated at the point of initial allocation. Calculation of the left hand sides of (12) and (13) at the same point gives H_0. From equation (14) we deduce theoretical differences $(\bar{n} - n_0, \bar{w} - w_0)$ between the initial and the equilibrium allocations. Using (15) we transform these differences into the differences of the shares Δs_i. The column of these differences will be referred to as New Economic Geography Factor ($NEGF$) and will be used in our empirical work as a potential predictor of the industry reallocation.

2.5 Dynamic extensions of the model

Up to now we considered only a static model with a static equilibrium and claimed that the actual observable dynamics could be explained by the difference between the equilibrium and the initial disequilibrium allocation. However, it could be reasonable to construct a dynamic version of the model describing the economy out of the static equilibrium. Only having this dynamic model we can analyze dynamic properties of the equilibrium such as global or local stability and a type of equilibrium (node, focus, saddle point or vortex). Indeed, short-run disequilibrium behavior can be substantially different from the global trend given by the found direction towards equilibrium.

A mechanical way to add dynamics to the model is to take (n, w) as phase space variables and postulate that dynamics of the economy is driven by imbalance between the current state (n_r, w_r) and the equilibrium $(\bar{n}_r, \bar{w}_r)$:

$$\begin{aligned} \dot{n}_r &= -\delta_n (n_r - \bar{n}_r), \\ \dot{w}_r &= -\delta_w (w_r - \bar{w}_r), \end{aligned} \tag{18}$$

where $\delta_n > 0$, $\delta_w > 0$ are parameters which control the rate of convergence. This system is equivalent to the assumption that the wages and the number of firms cannot change immediately, however, they tend to converge to the equilibrium. By construction, in this case the equilibrium is stable and small deviations from it will induce reverting dynamics. Note that predictions following from this simple dynamic model coincide with the predictions of the static model.

Another (more natural and less trivial) way to introduce dynamics is to assume that if there are n firms in the given region which produce n different goods and each firm earns positive profit π (this profit is the same for all firms due to similarity of production technology and symmetry in consumer preferences), then there are incentives for a new firm to enter the market with its own good and also get positive profit. Also it is reasonable to suppose that the process of entry is not instantaneous but takes time and the number of firms entering in each unit period of time is proportional to the available profit π. In this case the dynamics is given by $\dot{n}_r = \delta_n \pi_r$, where δ_n is a coefficient of firms inertia. If profit is negative, firms exit the market. Notably in this case the dynamics of the system is different from the trivial dynamics discussed above, but the equilibria are obviously the same. Full dynamic system has the following form:

$$\frac{l_r^m}{Fn_r} - 1 - \sum_{s=1}^{R} \frac{\mu_m \left(L_s + \frac{1}{\sigma_m - 1}\left(l_s^m - \sigma_m F n_s\right)\right)}{\sum_{q=1}^{R} F n_q \left(\frac{w_q}{\rho_m} T_{qs}\right)^{1-\sigma_m}} \left(\frac{w_r}{\rho_m} T_{rs}\right)^{-\sigma_m} T_{rs} = 0, \tag{19}$$

$$l_r^m - L_r(\mu_m + \mu_a) + \sum_{s=1}^{R} \frac{\mu_a w_s \left(L_s + \frac{1}{\sigma_m - 1}\left(l_s^m - \sigma_m F n_s\right)\right)}{\sum_{q=1}^{R} (w_q T_{qs})^{1-\sigma_a}} (w_r T_{rs})^{-\sigma_a} T_{rs} = 0, \tag{20}$$

$$\dot{n}_r = \frac{F \delta_n w_r}{\sigma_m - 1} \left(\frac{l_r^m}{F n_r} - \sigma_m\right). \tag{21}$$

Dynamic variables here are l_r^m, w_r and n_r. The variable q_r^m was excluded making use of the definition of regional industrial employment $l_r^m = n_r(F + c q_r^m)$. Deriving this system we also used equation (8) and the augmented form of income $Y_r = w_r L_r + n_r \pi_r$. Notably, parameter c drops out of this equation (just as out of the equations determining the static equilibrium) and parameter F enters only in combination with n_r (also δ_n should be rescaled by F).

Having equations (19), (20), (21) we can calculate $\dot{l}_r^m$ and after that the direction of changes in the distribution of industrial employment across regions in linear approximation:

$$\Delta s_i = \Delta \frac{l_i^m}{\sum_{r=1}^{R} l_r^m} \approx \frac{\dot{l}_i^m \sum_{r=1}^{R} l_r^m - l_i^m \sum_{r=1}^{R} \dot{l}_r^m}{\left(\sum_{r=1}^{R} l_r^m\right)^2}$$

It is worth to note that there is a significant difference between the implications of the static equations (12), (13) and the dynamic system equations (19), (20), (21). In the static equilibrium l_r^m and n_r are proportional to each other and in the calculation of the shares of industrial employment they can be used interchangeably. Moreover, the static equations can be written in terms of *shares* of industrial employment and for our analysis we do not need *absolute* values of employment. However, this is not true in the case of dynamic system. Though we can exclude Fn_r from the system (19), (20), (21) we cannot reformulate it consistently in terms of the shares of industrial employment and for calculation of Δs_i we need the values of l_r^m. Thus in this case there is no analog to the normalization procedure used in the static case.

The next problem with the suggested dynamic extension of the model is that it is not easy to get unambiguous predictions on stability of its equilibrium. The reason for this is that we do not know the exact location of the equilibrium but can estimate only the direction from the initial point towards the equilibrium. However, in order to analyze stability we need linearization of the system at the equilibrium point. The only approach that could potentially help is to assume that the equilibrium is close to the initial point and using the continuity of the Jacobian matrix[3] of the system (19) - (21) to make a conjecture that real parts of the Jacobian matrix eigenvalues calculated at the initial point have the same sign as at the equilibrium point. Unfortunately, our analysis shows that even small deviations from the initial point lead to significant shifts in the eigenvalues and some of them change the sign. It means that the suggested approach does not work.

One more drawback of the suggested dynamic formulation is that it implicitly assumes that the only imbalance of the initial point is in the number of firms but not in the wages or distribution of industrial employment. This is equivalent to the assumption that if we plug real data for the initial year into equations (19) and (20) then the equations will be satisfied (probably with statistical error). This is a very strong assumption and calculations show that it is not valid in our case.

To summarize, in order to predict short run dynamics we need a number of additional assumptions which will allow us to specify disequilibrium behavior of the system. Thus we conclude that our initial model is unable to give solid short run predictions but instead should be used only for explanations of long-run changes.

[3] To avoid confusion it is worth to mention that this Jacobian matrix is essentially different from the Jacobian matrix H' introduced in the previous section.

3 Empirical results

Mainly, the purpose of our empirical work is twofold. On one hand, we answer the question whether the industry reallocation (measured in terms of industrial employment) indeed took place. On the other hand, we check whether the new economic geography factors (represented by the theoretically constructed $NEGF$) indeed have played a noticeable role in this process. We presume that the transition processes which deeply affected the economy as a whole could not leave the regional industrial employment intact. However, there could be three essentially different patterns of this influence. The first pattern is pronounced dynamics understood as a process of reallocation of employment that continues at least for several years and has a certain direction. The second pattern is a structural break that can be described as a significant and persistent change of allocation structure, but which, in contrast to the dynamical pattern takes a short period of time. The third potential pattern is represented by random fluctuations, i.e. there are observable changes of industry allocation but these changes are not persistent and do not lead to evolution in a certain direction.

The first part of our empirical analysis is devoted to distinguishing these patterns. We start with testing a null hypothesis that there was no pronounced dynamics and the regional shares of industrial employment randomly fluctuated. Unfortunately, we have a very limited number of time observations (a time series with 12 observations is really poor) and this complicates the procedure and makes it less statistically convincing. However, even under these limitations we manage to demonstrate that our data are not consistent with the hypothesis that there was no dynamical trend.

In the second part of our analysis we test the predictions of our model on empirical data. We consider regressions of real changes in the shares of regional industrial employment on the theoretically constructed factor $NEGF$ and a set of control variables. It appears that the $NEGF$ predicts the direction of dynamics quite well and this result is not sensitive to a number of changes in the model specifications.

As a supplement to the analysis of these two major issues we compare the actual and predicted dynamics of geographical concentration indexes.

3.1 Data

Let us start with description of our data. The main object of our analysis is the regional structure of industrial employment. The main geographical unit that we consider is a region (sub'ekt federazii). In total we have 78 observations per year, as we do not consider autonomous regions (avtonomnye okruga) separately from the regions to which they belong and exclude Ingushetiya and Chechnya.

We collected annual data on the regional industrial employment for years 1990-2001 from Goskomstat annual reports “Labor and Employment in Russia” for 1997 and 2001. To our knowledge, the corresponding data for previous years were not officially published by Goskomstat. However, Goskomstat collected these data and they are available. We used data for 1985 from Horrigan (1992). Since in our work we analyze dynamics of the regional structure of industrial employment, absolute values of the regional industrial employment are not relevant for us. Thus, making use of the collected data y_{it} we construct the shares as

$$s_{it} = \frac{y_{it}}{\sum_{j=1}^{78} y_{jt}}$$

The obtained panel of shares s_{it} is the main object in our empirical exercises.

The described data are sufficient for testing the dynamics. However the analysis of the reallocation determinants requires more information. In particular, to be able to make our model testable we should specify a correspondence between model variables and the real data.

1. The simplest relation is between the parameters L_i of the model and the actual population. For L_i we take real regional population in the starting year which is 1985. We take these data from the Goskomstat annual report “Regions of Russia”, 2001.

2. To build a proxy for the model variable w we take regional money income per capita provided in the Goskomstat annual report “Regions of Russia”, 2001 and apply the normalization procedure described in Section 2.4.

3. We model consumer's preferences with a Cobb-Douglas function. To calibrate the utility function, we need estimates for parameters μ_a, μ_m, and μ_s. As these parameters stand for income shares that a consumer spends on corresponding product groups (i.e., services, manufacturing and agricultural products in our model) we suggest to use shares

of the corresponding sectors in GNP (as we do not have intermediate manufacturing sectors in the model, it seems to be the most reasonable approach). Table 3 shows GNP structure in Russia and the ratios μ_a/μ_m. Remarkably, inspite of significant growth of service sector the relative preferences between manufactured and agricultural products remained approximately the same. Taking our theoretical result that equilibrium industry allocation is determined only by the ratio μ_a/μ_m we get that the growth of the service sector should not drastically affect the equilibrium industry allocation. To construct $NEGF$ we take the ratio μ_a/μ_m equal to 0.2. This value is suggested by Table 3.

4. In the model the variables T_{ij} were introduced as the costs of transportation of one unit of good between regions i and j. In reality these costs depend on the distance between regions, the nature of the product, mode of transportation (car, railroad, air) etc. We obviously cannot incorporate all these details in our stylized model. Thus we will assume that the transportation costs depend only on the distance d_{ij} between regions[4]: $T_{ij} = f(d_{ij})$. In our regression analysis we use three types of transportation cost functions. As a benchmark, we use the same function as in Hanson (1998): $T_{ij} = \exp(a \cdot d_{ij})$, where a is a positive constant. Performing sensitivity analysis we use two other functions. One of them is linear, the other is a rough approximation of railway tariffs and has the form $T_{ij} = 1 + ad_{ij}^2$, where a is a positive constant. Calculation of d_{ij} is not a straightforward procedure as well since the regions are not points and there are several ways to define distances between them. We define d_{ij} as a distance between regional capitals. In this case d_{ij} can be calculated using the geographic coordinates of the capitals.

In order to distinguish the effect of new economic geography factor from other potential determinants of industry reallocation we also use a number of control variables. A complete list of them is given in Table 5. To construct $tjan_i$ we use the average January temperature taken from the annual report "Regions of Russia", 2001. To construct the fuel dummy we employ data on the regional industrial structure, also provided in the annual report "Regions of Russia", 2001. We classify the region as fuel oriented (and consequently assign 1 to the corresponding entry of the fuel dummy) if the share of the

[4]Indeed, in reality distance is a major determinant of the transportation costs especially if we take into account that almost all interregional freight, except for that going through pipelines, is carried by railroads in Russia.

fuel industry in this region exceeds 40%. Totally, we have 8 regions classified as fuel oriented. Similarly we construct the export dummy. Namely, we divided all regions into three types: export-oriented ones with export dummy equal to 1, import-oriented ones with -1 and mixed ones with the dummy value 0. To reduce potential endogeneity we collected data on the regional export-import operations[5] for 8 years and assigned 1 to the regions which had had positive net export in all observed years (in total, there are 19 export regions). Similarly, we assigned -1 to the regions with negative net export at least in 6 of the observed years (there are 12 regions of this type). All other regions were considered as mixed and were assigned 0.[6] The data on the density of railroads is also taken from the annual report "Regions of Russia", 2001.

3.2 Testing for persistency of changes in industrial employment distribution

In this Section we describe the tests that we have applied to identify the presence of persistent changes in the regional distribution of industrial employment (further we will call these persistent changes spatial dynamics). Only having this dynamics (but not random fluctuations of the regional shares of industrial employment) we can pose a question about the determinants of industry reallocation. Thus the results of the presented tests for spatial dynamics should be considered as a justification for the regression analysis.

The main problem we encounter in analyzing spatial dynamics is a small number of time observations (considering the period of 1991-2000 we have only 10 points). To overcome this obstacle, we apply two different tests each of which gives some evidence in favor of spatial dynamics. The first of them is based on a Wald-type statistic and annual changes in the shares. The other analyzes behavior of cumulative changes of regional shares of industrial employment. We do not consider such a characteristic of distribution dynamics as beta-mobility since it concentrates on a specific question of convergence of a

[5]Here we include only international trade operations.

[6]This lack of symmetry in the definition of export-oriented and import-oriented regions is due to the fact that in most years the total export value exceeded the total import value. In spite of some arbitrariness in our definition we believe that most of our results are robust and will not be significantly different with other reasonable definitions of export-oriented regions.

given variable to a common level whereas we are interested in identification of dynamics of any kind. Similarly, we do not employ measures of relative mobility because it can be that initial ranking of the regions according to the share of industrial employment keeps preserved but there are systematic persistent changes of these shares, i.e. spatial dynamics.

To perform the first test employing a Wald-type statistic, we model the dynamics of employment shares as a linear trend (this specification is inspired by shortage of time observations which makes it impossible to model the time series structure of observations):

$$\log\left(\frac{s_{it}}{s_{it-1}}\right) = a_i + \epsilon_{it}. \tag{22}$$

Here s_{it} are regional shares of industrial employment, a_i are constants, ϵ_{it} denote errors which are assumed to be i.i.d. in time dimension, i is the index of the region and t is the year of observation. Moreover, we assume that $Var(\epsilon_{it}) = \sigma_i$, i.e. our observations are homoscedastic in time. Since the sum of all shares is equal to 1, we exclude one region from consideration.

In this specification the test for persistent dynamics is equivalent to the test for joint insignificance of the coefficients a_i. Namely, our null hypothesis is the absence of dynamics $H_0 : a_i = 0$ for all i. Note, that due to a small number of time observations the tests for significance of separate a_i have very low power. However, the test that all a_i are equal to zero is much more powerful since it involves more observations and this hypothesis imposes more restrictions.

First, we calculate $y_{it} = \log\left(\frac{s_{it}}{s_{it-1}}\right)$ and estimate a_i as

$$\hat{a}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}. \tag{23}$$

To test the hypothesis $H_0 : a_i = 0$ for all i we construct a Wald-type statistic

$$W = T\sum_{i=1}^{N}\frac{\hat{a}_i^2}{\hat{\sigma}_i^2}, \quad \text{where} \quad \hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}(y_{it} - \hat{a}_i)^2. \tag{24}$$

It can be shown that under null hypothesis $H_0 : a_i = 0$ for all i and $T \longrightarrow \infty$ the Wald statistic W is asymptotically distributed as $\chi^2(N)$.

There is a concern that high value of the Wald statistic and, consequently, rejection of H_0 can be caused by high value of only small number of a_i's. If this is the case, we cannot interpret it as evidence of spatial dynamics. To avoid this objection, we recalculate the statistic with one and two excluded observations and show that the result does not change significantly. To be more conclusive we exclude the regions with the highest individual t-statistics.[7] The obtained Wald statistics and the corresponding p-values are presented in Table 4, Panels A and B. We consider two subsamples of our sample to demonstrate that results of our test are significantly different for different periods of time. For the period 1990-1997 the Wald statistic is highly significant and this is not true for 1994-2001. Thus we can conclude that major shifts in industrial employment took place in the beginning of the 1990s and further we will consider only this period.

The described simple test has several drawbacks. First of all, the number of time observations is so small that asymptotic results are probably unreliable and the distribution of the Wald statistic may be far from $\chi^2(N)$ (potentially this could result in high p-value of the Wald statistic). Next, the construction of the Wald statistic assumes that observations for units i are independent. Taking into account that in our case shocks can be spatially correlated, the latter assumption can be violated and this destroys validity of the test. To overcome these problems, we have to derive results for finite sample and to control explicitly for cross-correlation effects. The details of this procedure are presented in Appendix.

Our results show that small number of time observations notably affects the distribution of the test statistic which appears to be rather far from its asymptotic counterpart. However, even after this correction our Wald statistic is significant at 5% level. The test for spatial correlation rejects the hypothesis of independence of spatial observations (see details in Appendix), so we correct our test statistic for spatial correlation. In Panel C of Table 4 we report two p-values: one is asymptotic and is obtained from $\chi^2(77)$ distribution, the other one is for a simulated finite sample distribution. We see that correction for spatial correlations does not change our results significantly. Thus we can conclude that in spite of a very limited sample we have got some evidence that increments of the regional industrial employment have statistically distinguishable trend.

The second test we apply should be able to identify spatial dynamics even if industry

[7]This exclusion of the regions with the highest t-statistic is not absolutely innocuous since it could introduce data mining bias.

reallocation took place only in a few number of years (say, in 2 or 3), i.e. there was a structural break instead of a constant trend (in this case it is also reasonable to pose the question about its determinants and direction). To take into consideration this possibility and to distinguish persistent changes in industrial employment from random fluctuations around some constant level we construct a panel of cumulative differences

$$z_{it} = \log s_{it} - \log s_{i0}, \qquad i = 1..N, \quad t = 1..T.$$

If the structural break took place then z_{it}, $t = 1..T$ should have a mean which is significantly different from zero. For each region i we construct the corresponding t-statistics. We yield that 64 out of 78 t-statistics are different from 0 at 10% significance level, 57 are different from 0 at 5% significance level and 44 are not zero at 1% significance level. If the shares fluctuated randomly we would get much fewer significant t-statistics (on average 8 at 10% level, 4 at 5% level, 1 at 1% level). This result is strong evidence supporting the existence of persistent changes (trend or structural break) in industry allocation.

To conclude, spatial allocation of industry defined in terms of the shares of the regional industrial employment has undergone substantial changes. Our results show that these changes took place mainly at the beginning of the 1990's and were persistent. In the next section we identify the determinants of these changes.

3.3 Regression results

In this section we analyze whether the constructed theoretically New Economic Geography Factor ($NEGF$) can predict observed changes in industry allocation that took place at the beginning of the 1990s. To this end we consider regressions of actual differences against the $NEGF$ and several control variables. Our basic dependent variable is differences in the industrial employment shares $\Delta s_i = s_i^1 - s_i^0$, where s_i^1 and s_i^0 are shares of regional industrial employment in region i in two different years which are taken as a starting and a final point. In our analysis we take 1985 as a starting year and 1995 as a final year, however, when we perform sensitivity test we try other options.

To construct the numerical values of $NEGF$ we use equations (14), (17) from the theoretical part. Most relations between model variables and observable variables are described in Section 3.1. However, we should specify unobservable values ρ_a and ρ_m and choose the functional form of the relation between distances and transportation costs. Calibration of parameters ρ_m and ρ_a is a subtle operation. Indeed, we definitely know only that they belong to the interval $(0, 1)$. However, in literature there are several examples of calibrations of these parameters and we can base our inferences on these results. For example, Hanson (1998) estimates parameters of the structural Krugman (1991) model and finds that ρ lies within the interval $(0.8, 0.95)$, while Midelfart-Knarvik et al. (2000) use the value of the preferences parameter equivalent to $\rho = 0.8$. To verify that knowing exact values of these parameters is not crucial for establishing our results we also do sensitivity analysis.

As a benchmark case we take $\rho_m = \rho_a = 0.8$. Also we assume that the relation between distances and the transportation cost is exponential with the coefficient $a = 0.001$: $T_{ij} = \exp(0.001 d_{ij})$. Our basic regression equation has the following form

$$\Delta s_i = \alpha + \beta \cdot NEGF_i + (\text{control variables})'_i \cdot \gamma + \epsilon_i. \tag{25}$$

The list of control variables used in our regressions is given in Table 5. In general, they can be divided in two groups: the variables employed in the construction of $NEGF$ (such as s_i^0, L_i, w_i^0) and additional variables that could help to control for other potential explanations of industry reallocation.

There are several reasons to add the first group of variables into our regression. First of all, it is quite natural to suggest that the differences Δs_i are proportional to the initial values s_i^0. Indeed, it could be that the equilibrium shares do not have substantial variation and all variation in our predictor $NEGF_i \approx \bar{s}_i - s_i^0$ is due to the variation of s_i^0. Moreover, one can suspect that the constructed $NEGF$ does not aggregate information on the population, distances, initial shares and initial wages, but simply reproduces one of these inputs which itself can predict the changes in shares. To control for that, we add as control variables not only the initial shares s_i^0, but also the initial wages w_i^0 and the regional population L_i. To control for the simplest distance-related factor we add the distance to Moscow $Mdist_i$.

Besides the variables already used in the construction of $NEGF$ we would like to control for some other factors that could have influenced the industrial allocation pattern and could diminish the predictive power of the New Economic Geography Factor. In particular, a significant role can be played by physical geography factors, such as climate and availability of natural resources. To control for them we consider the average January temperature $tjan_i$ as an indicator of the climate severity and the fuel dummy $fuel_i$. Indeed, there are arguments (e.g., see Mikhailova, 2003) that in the Soviet period excessive number of plants were allocated in the cold regions and thereby the cost of cold was underestimated. The market could start correcting this misallocation and this could explain the observed dynamics. Controlling for the average January temperature can help to distinguish the "cost of cold" explanation of dynamics from the "new economic geography" explanation.

Next, there are essential reasons for including the export dummy $export_i$ as a control variable (the detailed description of this variable is given in Section 3.1). Indeed, our model does not take into account export revenues, however, export oriented industries were comparatively successful in the 1990s. So we can expect that the presence of export-oriented industries in a region will increase its industrial employment share in comparison with other regions. To separate this industry specific effect from the "new economic geography" effect we need the export control variable.

Besides the described factors the industry reallocation might be caused by the infrastructure differences between regions inherited from the Soviet period and one might suggest that industry in the regions with more developed infrastructure grows faster and industrial employment shares of such regions increase. In general, it is rather difficult to capture these infrastructure differences concisely since they could have various forms. In our research we approximate the development of the regional infrastructure by the development of the transportation system. Namely, we expect that if the inherited infrastructure really affects the future industry development then the regions with highly developed transportation system will show on average increase in the industry employment share. To control for this we add the density of railroads $rails_i$ to the control variables.

The first piece of evidence that $NEGF$ is not a replica of model inputs, but is a substantially new variable comes from the correlation matrix reported in Table 7. Indeed, it is worth to note that correlations between $NEGF$ and other variables are low. This justifies the usage of $NEGF$ as an additional explanatory factor. From Table 7 it follows that the initial share of industrial employment and regional population are highly correlated. This is very reasonable and further we use only the initial shares of employment as a control variable to avoid multicollinearity.

Next we estimate (25) for different combinations of control variables and report the results in Table 8. The simplest regression is regression (1) which does not include any control variables. It is remarkable that the coefficient before $NEGF$ is positive and significant at 1% level and $NEGF$ alone explains about 20% of total variation. Next we add the control variables which were used in construction of $NEGF$ (s_i^0 and w_i^0). From these regressions we get one more piece of evidence that $NEGF$ is a new variable aggregating a lot of information. Indeed, the coefficient before $NEGF$ is significant and very close to the coefficient in regression (1). Consequently, the predictive power of $NEGF$ does not originate from variation of s_i^0 or w_i^0. Further, the coefficient before s_i^0 is significant and this was also expected. Moreover, it is negative and this looks very natural since the initial shares enter into the difference with a negative sign.

In specifications (4) - (8) we add by one other control variables to distinguish their effect from the effect of $NEGF$. Notably, the coefficient before $NEGF$ is stable, positive and highly significant in all regressions and new regressors do not destroy its predictive power. The coefficients before control variables are also remarkable. Some of them (w_i^0, $Mdist_i$, $tjan_i$) are statistically insignificant and this shows that the initial distribution of wages, the distance to Moscow and the average temperature of January by themselves do not explain the observed industry reallocation dynamics. Notably the t-statistics of other control variables ($fuel_i$, $export_i$, $rails_i$) added by one are high and these variables seem to have explanatory power.[8] However, when taken all together in regression (9) most of them become insignificant and the export dummy appears to be the only important (and still significant at 1% level) factor. Thus it is reasonable to suggest that the high

[8]One can suspect that the coefficients before these variables are very small and thus these variables can be neglected as explaining a very small part of total variation. Partially this is true and their contribution to R^2 is not large, however, the scale of the coefficients is mainly due to the units of measurements. The products of the coefficients before $fuel_i$, $export_i$ and $rails_i$ and the standard deviations of these variables are equal to $2.77 \cdot 10^{-4}$, $5.81 \cdot 10^{-4}$ and $-3.56 \cdot 10^{-4}$ correspondingly. This is comparable with the similar product for $NEGF$ which is equal to $8.76 \cdot 10^{-4}$. Thus the small values of the coefficients is not a reason to discard the corresponding variables.

significance of $fuel_i$ and $rails_i$ in regressions (6) and (8) could be caused by correlations between these variables and the export dummy. Indeed, from Table 7 we know that the export dummy is positively correlated with $fuel_i$ and negatively with $rails_i$ (correlation with $rails_i$ is quite small). Notably the coefficient before $fuel_i$ in the regression (6) is positive and its sign coincides with the coefficient sign before $export_i$, but the coefficient before $rails_i$ in regression (8) is negative and has the opposite sign to the coefficient before $export_i$. Moreover, the export dummy makes the largest contribution to R^2. All this is evidence in favor of the claim that the only important variable is the export dummy. To be even more conclusive we consider the regressions (see Table 8, regressions (10) and (11)) in which the export dummy is taken together with one of $fuel_i$ or $rails_i$. One can see that the variables $fuel_i$ and $rails_i$ taken together with the export dummy become much less significant but the export dummy is still significant at 1% level. This is one more evidence that among the control variables only the export dummy is a real determinant of the industry reallocation dynamics and the presence of export oriented industries in the region makes it more likely that the industrial employment share of this region increases. As a result we will keep only the export dummy as a control variable in the further analysis.

There are several potential caveats in the reported results of $NEGF$ significance. One of possible objections is that this result may not be robust with respect to the chosen final year. To discard this objection, we perform a sensitivity test taking sequentially 1994, 1995, 1996 as a final year and reproducing regression (7) from Table 8. Results of this test are reported in Table 9. From this table it can be seen that coefficients before $NEGF$ are significant at 1% level and quite stable, though they increase from 0,149 in 1994 to 0,252 in 1996. Regressors s_i^0 and $export_i$ are also significant at 5% level for all three years and have the expected signs.

When we construct $NEGF$ we take ρ_a and ρ_m with essential arbitrariness. Thus we should check that our results do not significantly depend on particular values of these parameters (at least in appropriate interval). To be more concise we report only the coefficient before $NEGF$ with its p-value for regression specification (7) form Table 8. Again, Table 10 shows that $NEGF$ keeps to be significant at 1% level for all considered values of ρ_a and ρ_m. The values of the coefficient before $NEGF$ are very close to each other for ρ_m equal to 0.75 and 0.80, but differ for $\rho_m = 0.85$. Thus, we can conclude that though absolute values of the coefficient before $NEGF$ vary with ρ_m it does not influence their significance and validity of our results.

One more subtle point of our analysis is the choice of the function that relates distances between regions and transportation costs. However, just as in the case with constants ρ_a and ρ_m exact form of this function is not important for establishing our result. To be more convincing we report our regressions for several different forms of this function. We find that our results are robust with respect to the choice of transportation cost function: Table 11 demonstrates that neither regression coefficients, nor p-values vary considerably with the type of cost function and its parameters.

To summarize, all reported results show that the predictive power of the theoretically constructed factor $NEGF$ is not artificial or fragile. Contrary, it is valid for a number of model specifications and this testifies in favor of important and unambiguous role of "new economic geography" factors in the process of industrial reallocation in Russia in the 1990s.

3.4 Dynamics of geographic concentration

Having established that the spatial allocation of industrial employment was changing over the period of transition it is natural to pose the question whether the observed dynamics changed geographic concentration of industrial employment.

There are several ways to define the index of geographic concentration. In analogy to the standard analysis of monopoly power we consider two types of indexes:

1. Herfindahl-type index, which is defined as $HHI = \sum_{i=1}^{R} s_i^2$, where s_i is a share of industrial employment in region i.

2. Concentration ratio CR_k based on a total share of k largest regions and defined as $CR_k = \sum_{i=1}^{k} s_{(i)}$, where $s_{(i)}$ is the i-th largest share.

First, consider the Herfindahl-type index. The actual values of this index for years 1985 - 2000 are presented in Figure 1. It is easy to see that except a slight downward jump in 1985-1990 from 0,0236 to 0,022 there were no essential movements of this index. This is an indication that reallocation of industrial employment did not seriously affect the concentration.

Now let us consider the question whether the direction of dynamics predicted by our model leads to substantial changes in concentration or not. The simplest way to answer it is to consider the equilibrium allocation and compare the concentration indexes in the theoretical equilibrium and in the initial state. However, instead of the equilibrium shares of employment we only have the vector $NEGF$ which gives a direction towards the equilibrium and the new concentration indexes inevitably depend on absolute value of this vector. To overcome this problem we consider the index for a range of potential absolute values of the directional vector (we will denote the corresponding factor as λ) keeping in mind that the most probable value is given by the regression coefficient before $NEGF$ (approximately $\lambda = 0.17$). Thus for the Herfindahl-type index we have

$$HHI_\lambda = \sum_{i=1}^{R} \left(s_i^0 + \lambda \cdot NEGF_i\right)^2 = HHI_0 + 2\lambda \sum_{i=1}^{R} s_i^0 \cdot NEGF_i + \lambda^2 \sum_{i=1}^{R} NEGF_i^2. \quad (26)$$

HHI_λ is a quadratic function of λ and its graph is depicted in Figure 2. It can be seen that the predicted dynamics does not lead to essential changes of HHI. Indeed, for $\lambda = 0.17$ we have only small decline. Qualitatively the direction of movement coincides with the jump of the empirical Herfindahl-type index, however, the predicted size of the effect is smaller. In any case, both changes in the real and theoretically constructed indexes are pretty small and we can say that in general the predictions of the changes in concentration given by the model do not contradict to the observed values.

The results for the geographic concentration ratio are very similar. For brevity we report only the results for CR_4, however, the results for other k that we tried are very similar. CR_4 for the observed employment is presented in Figure 3. Just as in the case of Herfindahl-type index we observe a drop in 1985-1990 and after that CR_4 is quite stable. Contrary to the Herfindahl-type index the predicted concentration ratio index CR_4^λ for $\lambda \in [0, 1]$ is linear[9] with respect to λ: $CR_4^\lambda = CR_4^0 + \lambda Q$ where Q is a sum of $NEGF$ components corresponding to four regions with the largest shares of industrial employment (these regions are Moscow, Moscow region, Sverdlovsk region and St. Peterburg). For the empirically constructed vector $NEGF$ we get $Q = -0.0063$. Thus our model predicts

[9] It is not linear for all values of λ but only piecewise linear since for large λ ranking of the regions might change.

decrease of CR_4, however, the absolute value of this decrease is less than the observed one.

To summarize, the empirical data show that in 1985-1999 there were no drastic changes in concentration of industrial employment, however, in this period it slightly decreased. This behavior is in agreement with the predictions of our model.

4 Discussion and Conclusion

In our research we investigate a role of "new economic geography" factors in spatial reallocation of industrial employment in Russia in the 1990s. To characterize this process, we use regional shares of industrial employment. The basic results of our research are as follows:

1. There is statistical evidence that there were systematic and persistent changes in the pattern of industry allocation in Russia in 1990-1997.

2. "New economic geography" indeed gives an explanation of geographic reallocation of Russian industry in 1985 - 1995. This result appears to be valid for several model specifications including different preference parameters and transportation cost functions as well as the choice of the final year. Also this result cannot be eliminated by controlling for other potential determinants of industry allocation.

3. Among all considered variables controlling for other than "new economic geography" determinants of allocation of industry only the export dummy appears to be significant: regions that are mostly export-oriented increased their shares in the industry employment.

4. Geographic concentration of industry in 1985-1999 slightly decreased and this is also in agreement with the predictions of our "new economic geography" model.

It is worth to make several comments on our results. First of all, we do not claim that we have identified all factors influencing the reallocation of industry. Indeed, this is an extremely complicated process with a number of determinants. Our results show only that the "new economic geography" mechanism does work and as a by-side product we learn that export orientation is also one of the important factors. However, in particular regions some other factors that we do not consider can be relevant (for example, industrial policy of the regional government, regional industry profile, shocks in particular branches of industry and so on). Unfortunately, it is impossible to take into account all factors of this type (and the number of observations imposes serious restrictions on their testing). One might suggest that some of them are correlated with the "new economic geography" factor. However, the remarkable property of the constructed predictor $NEGF$ is its explicit exogeneity. Loosely speaking, $NEGF$ is a specific alloy of the inputs of the model (distances between regions, population of regions, initial industrial shares and wages) and these inputs are mixed very well. Thus $NEGF$ is not sensitive to particular input variable but aggregates information about geographic structure of economy. That is why it is reasonable to expect that $NEGF$ can be only weakly correlated with most of other "pure" variables taken as controls in the regression and this is actually true for the considered variables.

Next, it is worth to give some comments on generality of our model. One can argue that constructing a model we do a number of artificial assumptions and introduce a number of functional specifications which could strongly affect the predictions of the model. Partially, it can be true. Indeed, to make the model tractable and to get closed equations describing the equilibrium we need a number of artificial assumptions (and this is a common feature of most "new economic geography" models). Most of these assumptions (such as one factor of production, CES production functions, "iceberg" form of transportation cost) are actually crucial for tractability. However, the used assumptions are less strict and less bounding then they could seem. Indeed, in our theoretical part we show that despite the fact that we introduce the parameters c and F of industrial production function we do not need them for calculating the equilibrium shares. The constants μ_m, μ_a and μ_s finally enter only as a ratio μ_a/μ_m. Moreover, we can add a proportionality factor to the production function of the agricultural and service sectors and these factors also will drop out from the equations determining the equilibrium. The only fact that matters is that these functions are CRS. In our empirical part we show that our results are not sensitive to exact values of unobserved parameters ρ_m and ρ_a, i.e. our results are valid if these parameters are in the reasonable intervals. Moreover, we demonstrate robustness to the choice of functional form of the relation between distance and transportation costs.

In our study of industry employment dynamics we learned that the most active period of industry reallocation is the beginning of the 1990's. At the end of the 1990's this dynamics seems to disappear. Unfortunately, having an extremely limited number of observations we cannot say whether the economy indeed formed the equilibrium distribution or we observe a trend in another direction. Only data for sufficiently long period of time can help to answer this question.

It should be stressed that our model is applicable only to the industry as a whole but not to particular branches of industry. In spite of the fact that there are interesting questions about reallocation dynamics and its determinants in particular branches of industry research in this area would require specific models which would account for factors important for a given industry (such as links with suppliers and customers, export potential, availability of natural resources etc.). In this case new economic geography factors can be dominated by other factors which are completely ignored by our model. As a result, one can hardly hope that $NEGF$ extracted from the model with three stylized sectors will be helpful for explaining the dynamics of particular industries.[10]

There are several directions in which our research can be developed. The first one is application of the same approach to analysis of reallocation of industry in other countries. For example, it is possible to consider Ukraine, which as Russia experienced transition to the market economy, or Germany, in which case the unification of former GDR and West Germany might also have led to observable shifts in a regional distribution of industry.

Having complete plant level data set it would be possible to consider not regional industrial employment but the plant level one. This would get more insights into dynamics of industrial employment and would give an opportunity to analyze not only the industry

[10]Our analysis of geographic reallocation of employment in the fabric industry supports this claim. In particular, we test for spatial dynamics in this industry and perform regression analysis to check whether $NEGF$ remains a significant determinant of spatial reallocation. Overall, we have found that there is some evidence of spatial dynamics in fabric production. We run regressions of changes in the shares of regional employment in fabric production against $NEGF$ and other control variables (the same as described above except for initial shares of industrial employment which were substituted by initial shares of employment in fabric production). $NEGF$ appeared to be statistically insignificant in all these regressions. This supports our claim that $NEGF$ constructed from the model of 3-sector economy cannot explain dynamics in particular industries.

as a whole, but separate branches of the industry. Unfortunately, the development in this direction is substantially limited by imperfections and incompleteness of available data.

Appendix

In this Appendix we describe some details of the test for persistency of regional reallocation of Russian industry presented in Section 3.2. Namely, we explain the procedure of controlling for the finite sample effects as well as for potential spatial correlation between observations.

To control for finite sample effects, we make an additional assumption that the errors ϵ_{it} are normally distributed for each i. This assumption can be statistically justified by Lilliefors test of the hypothesis that the sample has a normal distribution with unspecified mean and variance against the alternative that the sample was not generated by normal distribution (see Conover (1980) for details). In our sample Lilliefors test rejects the normality hypothesis at 5% significance level just for approximately 5% of regions. Thus, the null hypothesis cannot be rejected. If errors have normal distribution then the distribution of the Wald statistic is also specified under the null hypothesis for finite T and coincides with the distribution of the following random variable

$$W_0 = T \sum_{i=1}^{N} \frac{\bar{\epsilon}_i^2}{\bar{\sigma}_i^2} \sim \sum_{i=1}^{N} t_i(T-1)^2,$$

where

$$\bar{\epsilon}_i = \frac{1}{T} \sum_{t=1}^{T} \epsilon_{it}, \qquad \bar{\sigma}_i^2 = \frac{1}{T-1} \sum_{t=1}^{T} \left(\epsilon_{it} - \bar{\epsilon}_i\right)^2,$$

and $\{t_i(T-1)\}$ is a set of independent random variables with Student distribution. Note that the resulting distribution does not contain the variances of errors. Instead of deriving an analytical form of distribution of W_0, we use simulations to find quantiles and these quantiles are used in calculation of p-values of the Wald statistic. Namely, we consider 1000 realizations of W_0 and calculate empirical quantiles. They appear to be significantly different from the asymptotic ones.

It is more difficult to adjust our results to potential spatial correlation of the errors ϵ_i. Since in our case $T < N$, we do not have sufficient data for constructing empirical estimator of the correlation matrix (the corresponding estimated matrix is degenerate). As a result, we should explicitly impose appropriate structure on the correlation matrix to reduce the number of its unknown entries. The simplest way to do this is to assume that the only factor affecting the value of spatial correlation is whether the given regions have a common border or not. To check validity of this factor we consider the following regression

$$\rho_{ij} = b_1 + b_2\delta_{ij} + \nu_{ij}, \qquad i = 2...N, \quad j = 1...i-1.$$

Here ρ_{ij} are empiric correlations between ϵ_{it} and ϵ_{jt} (we assume that population correlation coefficients do not change in time), $\delta_{ij} = 1$ if the regions i and j have a common border and $\delta_{ij} = 0$ otherwise, ν_{ij} is an error. Note that we consider only the upper triangular part of the correlation matrix. Results of the estimation show that both b_1 and b_2 are significantly different from zero. This implies that spatial correlation indeed takes place in the suggested form and we have to correct our test statistic. To do this we take the estimated covariance matrix in the form

$$\Sigma_{ii} = \frac{1}{T-1}\sum_{t=1}^{T}(y_{it} - \hat{a}_i)^2,$$

$$\Sigma_{ij} = (\hat{b}_1 + \hat{b}_2\delta_{ij})\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(y_{it} - \hat{a}_i)^2}\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(y_{jt} - \hat{a}_j)^2}, \qquad i \neq j$$

and construct the Wald statistic as

$$W = T\hat{a}'\Sigma^{-1}\hat{a}.$$

Tables and Figures

Table 1. List of the variables used in the model

L_r	-	population in the region r
w_r	-	wage in the region r
Y_r	-	aggregate income in the region r
l_r^m, l_r^a, l_r^s	-	total employment in the industrial, agricultural and service sectors in the region r
q_r^m	-	industrial output of one firm in the region r
q_r^a, q_r^s	-	output of agricultural and service sectors in the region r
p_r^m, p_r^a	-	prices for the manufactured and agricultural goods of the region r in the region r
P_r^m, P_r^a	-	composite price indexes of the manufactured and agricultural goods in the region r
n_r	-	the number of manufacturing firms in the region r
T_{rs}	-	"iceberg form" transportation costs between regions r and s
F, c	-	parameters of the industrial production function
μ_m, μ_a, μ_s	-	parameters of Cobb-Douglas utility function
ρ_m, ρ_a	-	weights of each manufactured and agricultural variety in the composite manufactured and agricultural goods

Table 2. Comparative characteristics of the model sectors

	Industry	**Agriculture**	**Service**
Type of product	differentiated	differentiated	homogeneous
Production function	$l^m = F + cq^m$	$l^a = q^a$	$l^s = q^s$
Transportation costs	$T^m = T$	$T^a = T$	$T^s = \infty$
Market structure	monopolistic competition	competitive	competitive

Table 3. Structure of GNP in Russia, 1985-2001

	1985	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
μ_m	0.62	0.58	0.50	0.63	0.49	0.35	0.39	0.38	0.36	0.37	0.39	0.40	0.38
μ_a	0.13	0.16	0.12	0.09	0.07	0.07	0.08	0.08	0.07	0.07	0.08	0.06	0.07
μ_s	0.25	0.26	0.38	0.28	0.44	0.58	0.53	0.55	0.57	0.56	0.53	0.53	0.55
$\frac{\mu_a}{\mu_m}$	0.21	0.27	0.23	0.14	0.14	0.21	0.19	0.21	0.20	0.18	0.19	0.16	0.17

Source: "Russian Statistical Yearbook", Goskomstat, 1997, 2001

Table 4. Testing for persistency of changes in industrial employment distribution

	Panel A: Simple Wald test, 1990 - 1997		
	All observations	1 observation excluded	2 observations excluded
W-statistic	165.3244	136.8209	119.6423
p-value	$2.1245 \cdot 10^{-8}$	$2.3676 \cdot 10^{-5}$	$8.0766 \cdot 10^{-4}$
	Panel B: Simple Wald test, 1994 - 2001		
	All observations	1 observation excluded	2 observations excluded
W-statistic	83.2687	73.0204	60.7633
p-value	0.2927	0.5756	0.8830
	Panel C: Wald statistics corrected for spatial correlation		
	W-statistic	Asymptotic p-value	Finite sample p-value
	168.4613	$8.8307 \cdot 10^{-9}$	0.043

Table 5. List of control variables

s_i^0	-	regional share of industrial employment in the year 1985
L_i	-	regional population
w_i^0	-	regional wages in the year 1985
$Mdist_i$	-	distance to Moscow
$tjan_i$	-	average temperature of January
$fuel_i$	-	fuel dummy
$export_i$	-	export dummy
$rails_i$	-	density of railroads

Table 6. Summary statistics

	$NEGF_i$	s_i^0	w_i^0	$Mdist_i$	$tjan_i$	$fuel_i$	$export_i$	$rails_i$
Mean	0	0.013	1.114	1803.5	-11.5	0.103	0.090	164.3
Std. deviation	0.005	0.012	0.296	1881.3	8.5	0.305	0.628	128.32
Median	$-3 \cdot 10^{-4}$	0.010	1.030	1112.1	-8.1	0	0	136.5
Maximum	0.017	0.057	2.341	6768.6	0	1	1	583
Minimum	-0.015	$3 \cdot 10^{-4}$	0.614	0	-35.6	0	-1	0

Table 7. Correlation matrix of $NEGF$ and control variables

	$NEGF_i$	s_i^0	L_i	w_i^0	$Mdist_i$	$tjan_i$	$fuel_i$	$export_i$	$rails_i$
$NEGF_i$	1.000	-0.176	0.170	0.165	0.283	-0.141	0.050	-0.177	-0.085
s_i^0		1.000	0.935	-0.105	-0.315	0.152	-0.039	0.280	0.488
L_i			1.000	-0.098	-0.278	0.151	-0.011	0.214	0.475
w_i^0				1.000	0.579	-0.525	0.115	-0.121	-0.294
$Mdist_i$					1.000	-0.749	0.010	-0.017	-0.564
$tjan_i$						1.000	0.103	-0.118	0.547
$fuel_i$							1.000	0.155	-0.198
$export_i$								1.000	-0.032
$rails_i$									1.000

Table 8. OLS regressions of Δs_i against $NEGF_i$ and control variables

Δs_i	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$const$	0.000 (1.000)	0.001 (0.002)	0.001 (0.113)	0.001 (0.147)	0.001 (0.095)	0.001 (0.071)	0.001 (0.076)	0.002 (0.008)	0.001 (0.010)
$NEGF_i$	0.177 (0.006)	0.150 (0.002)	0.153 (0.001)	0.153 (0.003)	0.150 (0.002)	0.150 (0.001)	0.171 (0.000)	0.157 (0.001)	0.186 (0.000)
s_i^0		−0.063 (0.015)	−0.062 (0.016)	−0.063 (0.020)	−0.062 (0.017)	−0.062 (0.014)	−0.077 (0.001)	−0.042 (0.100)	−0.067 (0.009)
w_i^0			0.000 (0.624)	0.000 (0.754)	0.000 (0.368)	−0.001 (0.362)	0.000 (0.806)	−0.001 (0.201)	0.000 (0.873)
$Mdist_i$				0.000 (0.922)					0.000 (0.083)
$tjan_i$					0.002 (0.237)				$-3{\cdot}10^{-5}$ (0.442)
$fuel_i$						0.002 (0.021)			0.001 (0.130)
$export_i$							0.001 (0.001)		0.001 (0.004)
$rails_i$								$-4{\cdot}10^{-6}$ (0.014)	$-3{\cdot}10^{-6}$ (0.087)
R^2	0.193	0.340	0.340	0.340	0.378	0.400	0.461	0.393	0.530
R^2_{adj}	0.182	0.320	0.313	0.303	0.344	0.367	0.432	0.360	0.474

p-values are in parentheses

Table 8. OLS regressions of Δs_i against $NEGF_i$ and control variables (continuation)

Δs_i	(10)	(11)
$const$	0.001 (0.002)	0.002 (0.007)
$NEGF_i$	0.167 (0.000)	0.172 (0.000)
s_i^0	−0.075 (0.001)	−0.062 (0.015)
w_i^0	0.000 (0.532)	0.000 (0.345)
$Mdist_i$		
$tjan_i$		
$fuel_i$	0.001 (0.100)	
$export_i$	0.001 (0.003)	0.001 (0.002)
$rails_i$		$-3{\cdot}10^{-6}$ (0.080)
R^2	0.495	0.485
R^2_{adj}	0.460	0.449

p-values are in parentheses

Table 9. Sensitivity test with respect to final year

Δs_i	1994	1995	1996
$const$	0.001 (0.067)	0.001 (0.076)	0.000 (0.743)
$NEGF_i$	0.149 (0.002)	0.171 (0.000)	0.252 (0.000)
s_i^0	−0.081 (0.000)	−0.077 (0.001)	−0.072 (0.013)
w_i^0	0.000 (0.845)	0.000 (0.806)	0.001 (0.565)
$export_i$	0.001 (0.001)	0.001 (0.001)	0.001 (0.002)
R^2	0.496	0.461	0.431
R^2_{adj}	0.465	0.432	0.400

p-values are in parentheses

Table 10. Sensitivity test with respect to the constants ρ_a and ρ_m

$\rho_m \backslash \rho_a$	0.75	0.8	0.85
0.75	0.183 (0.000)	0.178 (0.000)	0.165 (0.000)
0.8	0.171 (0.002)	0.171 (0.000)	0.148 (0.001)
0.85	0.092 (0.008)	0.122 (0.002)	0.120 (0.008)

p-values are in parentheses

Table 11. Sensitivity test with respect to the function relating distances and the transportation costs

	$T_{ij} = exp(ad_{ij})$			$T_{ij} = 1 + ad_{ij}$			$T_{ij} = 1 + (ad_{ij})^2$		
a	0.0005	0.001	0.002	0.0005	0.001	0.002	0.0005	0.001	0.002
$const$	0.001 (0.036)	0.001 (0.076)	0.001 (0.158)	0.001 (0.019)	0.002 (0.023)	0.001 (0.067)	0.001 (0.040)	0.001 (0.078)	0.001 (0.014)
$NEGF_i$	0.184 (0.000)	0.171 (0.000)	0.159 (0.002)	0.183 (0.000)	0.174 (0.000)	0.198 (0.000)	0.171 (0.000)	0.176 (0.000)	0.152 (0.002)
s_i^0	−0.080 (0.001)	−0.077 (0.001)	−0.074 (0.002)	−0.056 (0.000)	−0.084 (0.000)	−0.077 (0.001)	−0.084 (0.000)	−0.077 (0.001)	−0.075 (0.002)
w_i^0	0.000 (0.589)	0.000 (0.806)	0.000 (0.996)	0.000 (0.450)	−0.001 (0.401)	0.000 (0.870)	0.000 (0.560)	0.000 (0.842)	0.000 (0.995)
$export_i$	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)	0.001 (0.000)	0.001 (0.001)
R^2	0.486	0.461	0.434	0.488	0.475	0.495	0.477	0.471	0.435
R^2_{adj}	0.458	0.432	0.403	0.460	0.446	0.468	0.448	0.442	0.404

p-values are in parentheses

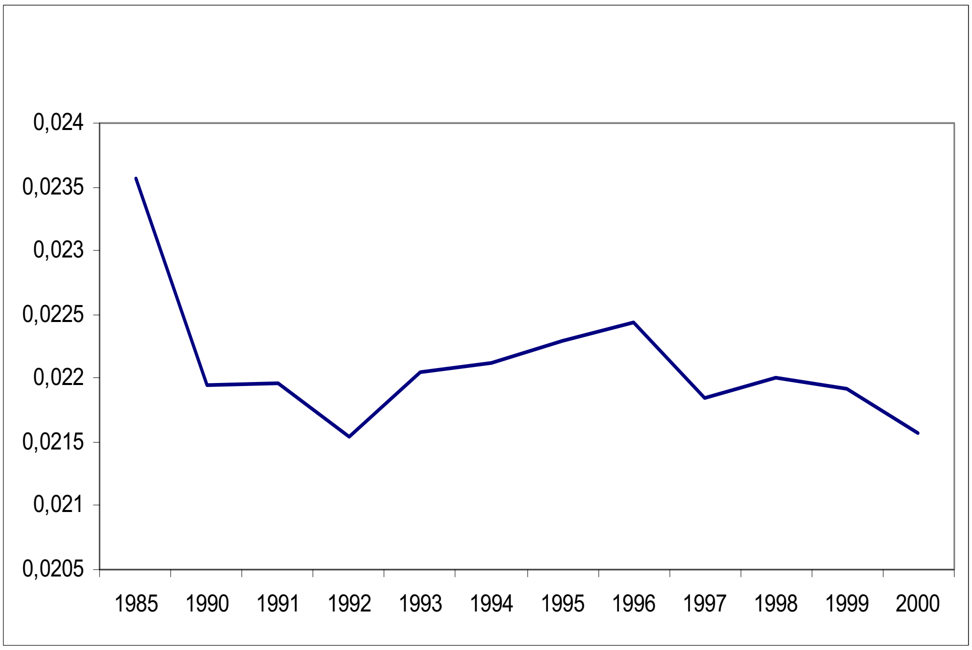

Figure 1. Empirical Herfindahl-type index of regional concentration of industrial employment

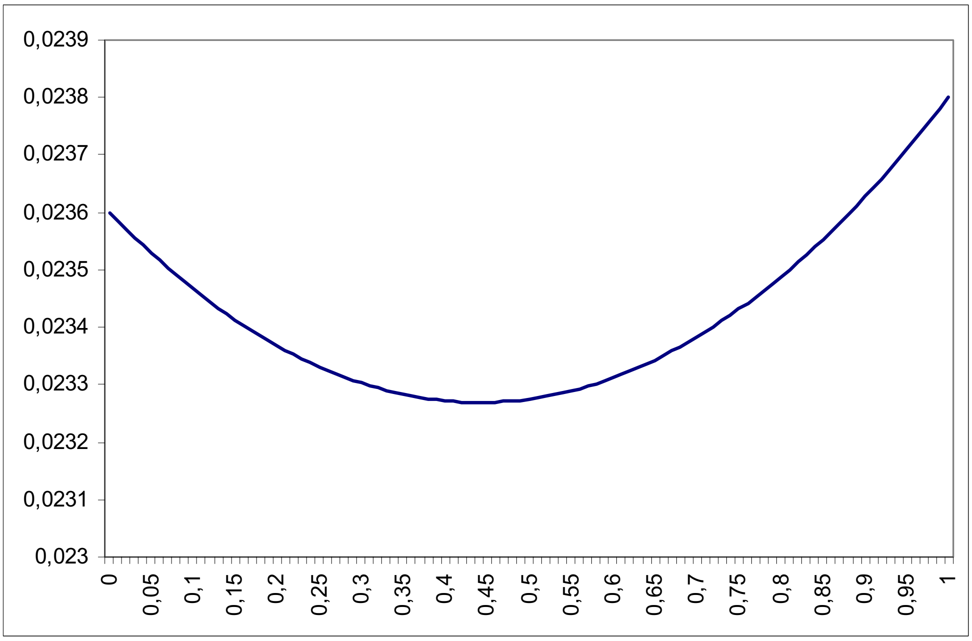

Figure 2. Theoretical Herfindahl-type index of regional concentration of industrial employment for $\lambda \in [0, 1]$

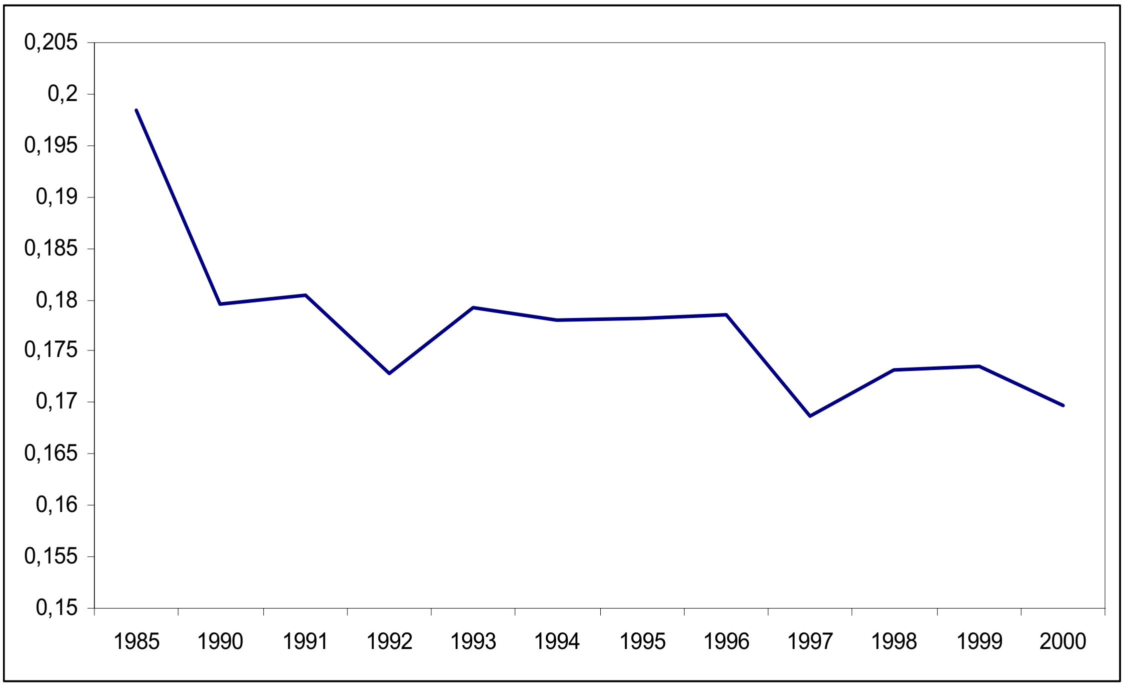

Figure 3. Empirical concentration ratio CR_4 for regional industrial employment

References

[1] Andrienko, Yu. and S. Guriev (2004) Determinants of interregional mobility in Russia, *The Economics of Transition* **12**(1), 1-27.

[2] Baldwin, R. E. (1999) Agglomeration and Endogenous Capital, *European Economic Review* **43**(2), 253-280.

[3] Baldwin, R. E. and R. Forslid (2000) The Core-Periphery Model and Endogenous Growth: Stabilizing and De-stabilising Integration, *Economica* **67**(267), 307-324.

[4] Brakman, S., H. Gerretsen, and M. Schramm (2001) New Economic Geography in Germany: Testing the Helpman-Hanson model, University of Groningen, Research Report No 01D46.

[5] Combes, P.-Ph. and M. Lafourcade (2001) Transport Costs Decline and Regional Inequalities: Evidence from France, CEPR DP No 2894.

[6] Conover, W. J. (1980) *Practical Nonparametric Statistics* (Wiley, New York).

[7] Davis, D. R. and D. E. Weinstein (1999) Economic Geography and Regional Production Structure: an Empirical Investigation, *European Economic Review* **43**(2), 379-407.

[8] Dixit, A. K. and J. E. Stiglitz (1977) Monopolistic Competition and Optimum Product Diversity, *American Economic Review* **67** (3), 297-308.

[9] Fujita, M., P. Krugman and A.J. Venables (1999) *The Spatial Economy* (MIT Press).

[10] Hanson, G. H. (1998) Market Potential, Increasing Returns, and Geographic Concentration, NBER WP No 6429.

[11] Horrigan, B. (1992) How Many People Worked in the Soviet Defence Industry? RFE/RL Report, v.1, No 33.

[12] Krugman, P. (1991) Increasing Returns and Economic Geography, *Journal of Political Economy* **99**(3), 483-499.

[13] Krugman, Paul, and A. J. Venables (1995) Globalization and the Inequality of Nations, *Quarterly Journal of Economics* **110**(4), 857-880.

[14] Martin, Ph. and G. I. P. Ottaviano (2001) Growth and Agglomeration, *International Economic Review* **42**(4), 947-968.

[15] Mikhailova, T. (2003) Essays on Russian Economic Geography: Measuring Spatial Inefficiency, Ph.D. Thesis.

[16] Midelfart-Knarvik, K. H., H. G. Overman, and A.J. Venables (2000) Comparative Advantage and Economic Geography: Estimating the Location of Production in the EU, CEPR DP No 2618.

[17] Ottaviano, G. I., T. Tabuchi, and J.-F. Thisse (2002) Agglomeration and Trade Revisited, *International Economic Review 43*(2), 409-436.

[18] Puga, D. and A. J. Venables (1999) Agglomeration and Economic Development: Import Substitution vs. Trade Liberalisation, *The Economic Journal* **109**(455), 292-311.

[19] Samuelson, P.A. (1954) The Transfer Problem and Transport cost, II: Analysis of Effects of Trade Impediments, *Economic Journal* **64**, 264-289.

[20] Venables, A.J. (1996) Equilibrium Locations of Vertically Intergrated Industries, *International Economic Review* **37**, 341-359.